Interactive comment on "Technical note: A simple approach for efficient collection of field reference data for calibrating remote sensing mapping of northern wetlands" *by* Magnus Gålfalk et al.

Anonymous Referee #2

Received and published: 16 December 2017

The issues this technical note seeks to remedy is truly an issue when working with spatial data and a hurdle for all researcher to overcome. The methods outlined in this research provide a low cost and easily accessible solution to not only researches but industry as well. Overall the paper is well written, and the methods seem useable. For a single researcher the methods provide a simple solution to a sometimes underestimated problem, especially those with limited time and resources. The even more important aspect to this research is that its simplicity and relatively inexpensive methods could lead to a database and data sharing. This is not possible with complex or expensive methods require resources and expert training and not to be underestimated. The equipment is both originally attainable on a budget but also common and

easily replaced if malfunctions or issues arise during a field campaign. The personnel needs are also cheaper and more available, making research and data collection more likely to occur. With a few edits outlined below this paper would be publishable and a help to benefit geospatial science as a whole. My first question is how this compares to the methods used currently. How long did it take to collect the 200 reference data points versus if they were collected in the "typical" fashion? (page 3) What is the computer processing time to create and classify the images and is that a major addition in time spent creating reference data? This system could potentially trade time in the field for time behind a computer time, not a bad option as processing time can be spread out over multiple analysts and a larger time frame. This methodology seems like it worked but did it? No discussion is made on how the results from the 200 plots performed and if they would help in classification. The level of accuracy achieved is an important metric for many and if this shows similar or a marginal decreases in accuracy with greatly reduced time or cost in data collection it would be well received. The risk of reduced accuracy may not be worth researcher changing the methods they currently use and some discussion on expected results of this new method should be included. The clarification of this topic would explain how data is integrated into classification of land cover and how this method compares to the current methods in terms of time savings and accuracy estimates? The section on distortion models is important for the research and could throw off all results for anyone using it. (page 4) What model was used and why was it used, enough explanation should be included for researchers to understand the difference distortion models would make on results and the sensitivity of different models. This is important especially if comparisons between research groups or sharing of data is going to occur. Calling validation "ground truth" should not be done, even if it's explained. The fact that it was explained means that there is an understanding that it's a bad term, call it reference data and remove any reference to "ground truth." The next sentence goes on to state 100% accurate reference data is not possible, reference data should be 100% accurate within the margin of possible error, it is not however 100% guaranteed that it represents the population or a large

study area, nor is it 100% guaranteed due to geolocation error that it is correctly located in your image compared to the true ground location. Clarification on this topic by the author would be helpful. Higher vegetation causing problems is made (page 7), but what is high vegetation and when does it start to deteriorate the data? Was high vegetation seen in the 200 plots created here and at what rate did tall vegetation cause problems? Knowing this would allow a decision on the applicability of this method to different research. My final comments on this technical note are about the writing mechanics. The sentences and paragraphs are well written, however in a few cases they start weak with; however, for example, to resolve this, etc. An effort should be made to start sentences with the primary subject of the sentence and tighten up some of the language and remove extra words seen throughout the paper. The acknowledgement heading is floating on page 10 line 15 as well. The additions and clarifications outlined above would make this an article worth publishing.